**Summary**

NetApp® SnapProtect™ management software is changing today's backup and recovery landscape. SnapProtect software combines simplified manageability, power, and flexibility for virtual environments with full support for enterprise database applications while providing virtually seamless integration with NetApp Snapshot® technology for fast and efficient backup operations. In addition, SnapProtect software integrates with NetApp SnapVault® and SnapMirror® software with support for content-based cataloging and movement to tape-based media. This document is an introduction to the SnapProtect solution. It provides an overview of the technology, and describes some of the basic configuration steps required to get started.
Figure 2) Unified Management 7-mode, SPOS, and clustered Data ONTAP. ........................................5
Figure 3) Distinguishing between node, HA pair, cluster, and Storage Virtual Machine. ......................8
Figure 4) NetApp NAS NDMP iDA ........................................................................................................9
Figure 5) Windows File System iDA .....................................................................................................10
Figure 6) OnCommand (formerly DFM) handling of storage provisioning and replication. ....................11
Figure 7) OnCommand 6.0 clustered Data ONTAP default provisioning policies. ................................12
Figure 8) SnapProtect Example Workflow: Vault -> Mirror -> Tape. ....................................................17
Figure 9) SnapProtect for Open Systems Technical Illustration ..........................................................19
Figure 10) SnapProtect for clustered Data ONTAP Workflow ...............................................................21
Figure 11) Mirroring NAS volumes. .....................................................................................................23
Figure 12) Vaulting NAS qtrees - fanin .................................................................................................24
Figure 13) Vaulting NAS volumes and qtrees - fanin. ...........................................................................24
Figure 14) Poor storage layout with multiple clients with LUN data on a common volume. ..................25
Figure 15) Better storage layout with multiple clients having LUN data on separate volumes ..............25
Figure 16) Efficient storage layout of single client with LUN data on a common volume. ....................26
Figure 17) Flexibility of replication combinations ................................................................................28
Figure 18) Example schedules and retention at specific times ..............................................................29
Figure 19) Application support, iDA, and restore granularity .................................................................32
Figure 20) Server Agent layers ............................................................................................................33
Figure 21) Datastores in separate subclients based on differing storage policies. ...............................34
Figure 22) Datastores in the same subclient, fan-in .............................................................................34
1 Introduction

NetApp is an industry leader in array-based data protection. The efficiencies of Snapshot copy technology and data replication have changed the way we look at backup and recovery and disaster recovery strategies. The need to achieve higher SLAs and to meet backup windows is a constant challenge, given the explosion of data that enterprises are dealing with today. Data center consolidation through virtualization has also created challenges around data protection. Disk-to-disk data protection solutions are becoming more widely accepted for both backup and recovery and disaster recovery strategies. NetApp data protection solutions offer speed and flexibility while reducing storage capacity requirements using efficient array-based technologies. The result is a simplified approach that reduces costs and administrative effort.

NetApp SnapProtect management software offers enterprise-class management for backup and recovery in the data center. The SnapProtect software manages Snapshot copies on NetApp primary storage and replication to secondary and tertiary storage, as well as tape creation. Irrespective of whether you are protecting NetApp application data, file data for NAS, file data in LUNs, or data in virtualized environments, the SnapProtect solution provides the management, the storage provisioning, the cataloging, and the granular recoverability required for seamless operation. Figure 1 shows the basic flow of the SnapProtect software.

Figure 1) SnapProtect software overview.

SnapProtect software can be used to protect the following applications hosted on NetApp primary storage:

- Microsoft Exchange (including DAG configs)
- Microsoft SQL Server® (Windows)
- Microsoft Office SharePoint® Server (Windows)
- Oracle® (Windows/UNIX/Linux) including RAC
- DB2 (UNIX and Linux)
- SAP® for Oracle (UNIX and Linux)
- Lotus Domino (Windows)

In addition, the SnapProtect solution supports the following virtualization products:
- VMware® vSphere™
- Microsoft Hyper-V™

For a complete list of supported platforms and product versions, refer to the Interoperability Matrix Tool (Resources section).

Figure 2) Unified Management 7-mode, SPOS, and clustered Data ONTAP.
1.1 Terminology

The following SnapProtect components work together to create a full solution.

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CommCell</td>
<td>A single instance of a SnapProtect environment.</td>
</tr>
<tr>
<td>CommServe</td>
<td>The master server in a SnapProtect environment. This server uses a Microsoft SQL Server database and therefore must be a Microsoft Windows system (Windows Server®).</td>
</tr>
<tr>
<td>Media agent</td>
<td>A media server in a SnapProtect environment (helper server to CommServe). Media agents have broad operating system support, including Windows, Linux, and UNIX options.</td>
</tr>
<tr>
<td>CommCell Console</td>
<td>The SnapProtect management interface.</td>
</tr>
<tr>
<td>iDataAgent (iDA)</td>
<td>Agents that control data consistency during backup operations.</td>
</tr>
<tr>
<td>Clients</td>
<td>Hosts running iDataAgents for which data is protected.</td>
</tr>
<tr>
<td>Backup Set</td>
<td>A layer of management within iDataAgents for grouping subclients.</td>
</tr>
<tr>
<td>Subclient</td>
<td>A layer of management within a Backup Set. A client can have multiple subclients, each of which can be associated with different source data.</td>
</tr>
<tr>
<td>Disk Library</td>
<td>A storage resource with an associated mount path that is used in the SnapProtect solution to store index information backups.</td>
</tr>
<tr>
<td>Storage Policy</td>
<td>A logical object through which a subclient is protected. The storage policy defines how data is backed up and replicated as well as retention requirements.</td>
</tr>
<tr>
<td>NetApp primary</td>
<td>The production NetApp storage array.</td>
</tr>
<tr>
<td>Component</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>NetApp secondary</td>
<td>The secondary NetApp storage array used as a destination for replication.</td>
</tr>
<tr>
<td>NetApp tertiary</td>
<td>A third NetApp storage array used for replicating previously replicated data.</td>
</tr>
<tr>
<td>Snapshot copy</td>
<td>A NetApp array-based point-in-time copy used for recovering data.</td>
</tr>
<tr>
<td>SnapVault</td>
<td>A NetApp replication technology used for backup and recovery. In the SnapProtect solution, a “vault” copy uses SnapVault.</td>
</tr>
<tr>
<td>SnapMirror</td>
<td>A NetApp replication technology used for disaster recovery. In the SnapProtect solution, a “mirror” copy uses SnapMirror.</td>
</tr>
</tbody>
</table>

2 Clustered Data ONTAP Overview

2.1 Basic Terminology

- **Cluster**: Consists of one or more nodes that are interconnected and managed as a single system.
- **Cluster interconnect**: A dedicated high-speed, low-latency, private network used for communication and replication between nodes in the same cluster.
- **Clustered Data ONTAP**: The Data ONTAP operating mode that supports interconnection of nodes into a cluster.
- **Data network**: The network used by clients to access data.
- **HA interconnect**: The dedicated interconnect between two nodes in one high-availability (HA) pair.
- **HA pair**: Two nodes configured in a pair for HA.
- **Ifgrp**: A collection of physical ports combined to create one logical port used for link aggregation; an integration group.
- **LIF**: A logical interface that is assigned an IP address that provides an Ethernet access point to a particular node in the cluster.
- **Intercluster LIF**: A LIF used only for intercluster replication, assigned only to a node.
- **Intercluster network**: The network used for communication and replication between different clusters.
- **Management network**: The network used for administration of the cluster, Storage Virtual Machine, and nodes.
- **Node**: A single NetApp controller, or one of a high-availability pair.
- **Port**: A physical port, such as e0e or e0f, or a logical port such as a virtual LAN (VLAN) or an interface group (ifgrp).
- **Storage Virtual Machine**: A logical storage server that provides data access to LUNs and/or a network-attached storage (NAS) namespace from one or more logical interfaces (LIFs).
2.2 Overview

Clustered Data ONTAP 8.1 is the first release that allows replication between different clusters, providing cross-site DR capabilities. New capabilities have been introduced in clustered Data ONTAP 8.1 for the following purposes:

- **Cluster peering**: The act of connecting two clusters to allow replication to occur between them.
- **Intercluster logical interfaces**: Logical network interfaces used for intercluster communication.
- **Intercluster ports**: Ports dedicated to intercluster replication.

Clusters must be joined in a peer relationship before replication between different clusters is possible. Cluster peering is a one-time operation that must be performed by the cluster administrators. The following steps are required to allow replication between different clusters:

1. Have a clear understanding of cluster peering.
2. Determine whether or not to share ports for data access and intercluster replication.
3. Designate ports for intercluster replication if dedicating ports.
4. Create intercluster LIFs on each node in the clusters.
5. Peer the clusters together.

Cluster peering must be performed because this defines the network on which all replication between different clusters occurs. Additionally, starting in clustered Data ONTAP 8.2, Storage Virtual Machines must be joined in a peer relationship before replication between different Storage Virtual Machines is possible. These tasks can all be completed via the OnCommand Unified Manager interface.

Storage Virtual Machine peering is the act of connecting two Storage Virtual Machines to allow replication to occur between them (starting in clustered Data ONTAP 8.2). In clustered Data ONTAP 8.1 any Storage Virtual Machine could replicate data to any other Storage Virtual Machine in the same cluster or any cluster peer. Control of replication security could only be maintained at a clusterwide level. Starting in clustered Data ONTAP 8.2, more granularity in replication security is provided. Replication permission must be defined by peering Storage Virtual Machines together.

Before creating any relationships between a pair of Storage Virtual Machines, you must have a Storage Virtual Machine peer relationship between the pair of Storage Virtual Machines. These Storage Virtual Machines can be local (intracluster) or remote (intercluster). Storage Virtual Machine peering is a permission-based mechanism and is a one-time operation that must be performed by the cluster
administrators. The following steps are required to configure replication between different Storage Virtual Machines:

1. Have a clear understanding of Storage Virtual Machine peering.
2. Peer the Storage Virtual Machines together.
3. Create NetApp SnapMirror/SnapVault relationships between different Storage Virtual Machines.

Best practice is to name a Storage Virtual Machine with a unique fully qualified domain name (FQDN): for example, dataNasSVM.HQ or mirrorDisasRecoSVM.Offsite. Storage Virtual Machine peering requires unique Storage Virtual Machine names, and using FQDN naming style makes it much easier to make sure of uniqueness.

3 Technical Overview

This section covers the technical details of the SnapProtect software and how the components work together.

3.1 Basic Functions

The SnapProtect solution delivers several basic functionalities to create a simplified user experience, including the following:

- Snapshot copy creation
- Cataloging and indexing
- Storage provisioning
- Data replication using SnapVault and/or SnapMirror
- Data movement to tape

Snapshot Copy Creation

The SnapProtect software creates Snapshot copies on the NetApp primary storage as its first backup copy. This is important because Snapshot technology allows backups to complete very quickly. Primary Snapshot copy creation is handled differently for different types of data. For NAS data, the NetApp primary system is treated as a client with an associated iDataAgent (iDA) called “NetApp NAS NDMP.” Subclients within the iDA are configured and associated with the NetApp data that requires protection. When a backup for the subclient runs, NetApp Snapshot copies are created for the volumes in that subclient. Figure 4 shows how this structure looks for a NetApp primary and its iDA.

Figure 4) NetApp NAS NDMP iDA.

For LUN data hosted on NetApp primary storage, the host accessing the data is treated as the client. The attached drive on the client is associated with a subclient within the File System iDA or the associated application iDA. On Windows clients, the iDA calls Microsoft Volume Shadow Copy Service (VSS) to
make sure that the data within the file system is consistent. With application-integrated Snapshot copies, the application agent calls VSS (Windows) or places the database in hot-backup mode (UNIX or Linux) for backup consistency. Then the Snapshot copy is created on the NetApp primary system for the volume containing the LUN. Figure 5 shows how this structure looks for a Windows client and its File System iDA.

**Figure 5) Windows File System iDA.**

---

**Cataloging and Indexing**

The ability to index the contents of a backup is a core value of the SnapProtect solution. For basic NAS data, the contents of the Snapshot copies that are created by the SnapProtect software are indexed directly. For LUN data, LUN clones are created and used for indexing. For LUN data, the contents inside the LUN are indexed. Specific proxy servers can also be assigned to handle indexing.

This indexing process can also be **deferred** until after the backup has successfully transferred to the secondary; this is a great feature allowing user to decouple backup and indexing schedules and have further control over processing cycle utilization. This option is a great feature to have when there are tight backup windows on Production controllers, a lot of files, and you simply want to complete the backup job as quickly as possible; you can specify in your policy to use the primary or secondary (vault or mirror snapshot copy) to run the indexing operation on. As usual, no matter whenever and wherever you choose to index your backups, once that indexing operation completes, you have a full catalog to browse and search through.

The indexes are stored in disk libraries. NetApp recommends creating disk libraries with paths that point to NetApp primary storage.

---

**Storage Provisioning**

Storage provisioning is required whenever Snapshot copies on NetApp primary storage need to be replicated to NetApp secondary and tertiary storage. Before replication can be established, the secondary system must have the appropriate volumes in place, along with the correct volume settings. SnapProtect software takes care of this by using the provisioning services of NetApp OnCommand Unified Manager (OCUM).

OnCommand uses policy-based rules that define how storage should be provisioned under different circumstances. Where the storage is provisioned is also flexible in that it uses pools of storage called resource pools—OnCommand containers that point to one or more aggregates within a NetApp system. The storage administrator only needs to provide the backup administration resource pools to use. The backup administrator simply directs the SnapProtect software to use the resource pools and provisioning policies for replication purposes. The appropriate storage is provisioned automatically.

Figure 6 shows how OnCommand handles provisioning and replication. When a new replication relationship is configured in SnapProtect management software, this information is passed to OnCommand, which creates the datasets needed to manage the replication requirements. OnCommand then provisions the necessary volumes, using the resource pools, and provisioning policies that were assigned in the SnapProtect configuration.
The resource pools must be created manually by using the OnCommand Unified Manager interface. The SnapProtect software then discovers the resource pools, and makes them available in the CommServe console for selection.

It is not necessary to create provisioning policies manually unless custom policies are needed; the SnapProtect software in relation with OCUM 5.x has a set of preconfigured provisioning policies (See Table 1).

Table 1) Preconfigured provisioning policies for OnCommand 5.x.

<table>
<thead>
<tr>
<th>Policy Name</th>
<th>Availability</th>
<th>Deduplication</th>
<th>Space Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td>SnapProtect_RAID-DP</td>
<td>RAID-DP</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SnapProtect_Dedupe</td>
<td>RAID-DP</td>
<td>On-Demand</td>
<td>80%, 90%</td>
</tr>
<tr>
<td>SnapProtect_Mirror_Destination</td>
<td>RAID-DP</td>
<td>No</td>
<td>80%, 90%</td>
</tr>
</tbody>
</table>

It is considered a best practice to use NetApp RAID-DP® for resiliency. Therefore, all of the preconfigured provisioning policies enable RAID-DP.

Deduplication can be enabled or disabled for vault copies independently of the deduplication setting on the primary data volume. Deduplication for vault copies uses the On-Demand setting, and deduplication runs automatically as vault copy jobs complete. To enable deduplication for vault secondary storage, the SnapProtect_Dedupe provisioning policy can be used. For vaulting that does not require deduplication on secondary storage, the SnapProtect_RAID-DP policy can be used.

Mirror copies inherit the deduplication settings of the primary data volume. Therefore, deduplication is disabled in the SnapProtect_Mirror_Destination provisioning policy. If the primary volume has deduplication enabled, the mirror copy volume is also a deduplicated volume.

Space thresholds represent the “nearly full threshold” and “full threshold” properties used when provisioning storage.
Custom provisioning policies are also available in OCUM 5.x. If there is a need for settings apart from the preconfigured policies, new provisioning policies must be created by using the NetApp Management Console. The SnapProtect software will automatically discover the custom provisioning policies. Custom provisioning policies must begin with a prefix of “SnapProtect _” to appear in SnapProtect.

In addition to OnCommand 5.x for 7-mode, OnCommand 6.0 provides comparable functionality, but is specific to clustered Data ONTAP environments. SnapProtect requires an OnCommand 5.x server for 7-mode and SnapProtect for Open Systems (SPOS) data, but also requires an OnCommand 6.0 server to handle clustered Data ONTAP data if applicable. One difference to note about SnapProtect’s support for OnCommand 6.0 is that, much like Table 1 above, OnCommand has default provisioning policies. Instead of the three that OnCommand 5.x has, OnCommand 6.0 has four default provisioning policies. Please see Figure 6 below for an illustration of these intuitive OnCommand 6.0 defaults.

**Figure 7** OnCommand 6.0 clustered Data ONTAP default provisioning policies.

4 Default Provisioning Policies

1. SnapMirror Destination
2. SnapVault Destination
3. SnapVault Destination with Deduplication
4. SnapVault Destination with Deduplication and Compression
Table 2) Preconfigured provisioning policies for OnCommand 6.x.

<table>
<thead>
<tr>
<th>Policy Name</th>
<th>Availability</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SnapMirror Destination</td>
<td>RAID-DP</td>
<td>Policy on the secondary volume used for mirror replication</td>
</tr>
<tr>
<td>SnapVault Destination</td>
<td>RAID-DP</td>
<td>Policy on the secondary volume used for vault replication</td>
</tr>
<tr>
<td>SnapVault Destination with Deduplication</td>
<td>RAID-DP</td>
<td>Policy with deduplication enabled on the secondary volume used for vault replication</td>
</tr>
<tr>
<td>SnapVault Destination with Deduplication and Compression</td>
<td>RAID-DP</td>
<td>Policy with deduplication and compression enabled on the secondary volume used for vault replication</td>
</tr>
</tbody>
</table>

Note: SnapProtect does not allow creating new custom Provisioning Policies for clustered Data ONTAP.

3.2 Data Replication Using SnapVault and/or SnapMirror

As mentioned earlier, there are two types of replication operations that can be configured using the SnapProtect solution — vault copies and mirror copies. It is important to understand the difference between these two types of replication operations.

A vault copy uses NetApp SnapVault, and a mirror copy uses NetApp asynchronous volume SnapMirror. SnapVault and SnapMirror replicate data in a similar way in that they both replicate only the blocks that have changed since the last replication operation. A full copy of the data is made only once. After the initial full copy, block-based incremental backups are made. This approach to disk-to-disk data protection offers speed, network efficiency, and storage capacity savings.

A vault copy has a certain independence from the primary data, allowing retention levels between the primary copy and the secondary copy to be different. For example, a vault copy allows longer-term retention than the primary copy.

A mirror copy depends more firmly on the primary data and is an exact mirror of the source volume and its Snapshot copies. A mirror copy cannot have an independent level of retention. Mirror copies are traditionally used in disaster recovery solutions because they allow failover to the secondary copy. Another use for mirror copies is to duplicate a vault copy to a remote location.

SnapProtect allows several combinations of vaulting and mirroring, satisfying many disk-to-disk-to-tape requirements.

Cascading SnapMirror and SnapVault relationships are supported in clustered Data ONTAP 8.2. A SnapMirror secondary can be the source of a SnapVault relationship (backing up a DR mirror); or a SnapVault secondary can be the source of a SnapMirror relationship (protecting a backup). In the case of a SnapMirror to SnapVault relationship, it is not possible to specify which Snapshot copies are transferred to the SnapVault secondary; SnapVault always transfers the SnapMirror exported Snapshot copy, or the base copy of the SnapMirror relationship. This is similar to the “snapmirror base snapshot only” option in 7-Mode.
3.3 SnapVault for clustered Data ONTAP

SnapVault was rebuilt from the ground up for its debut in clustered Data ONTAP 8.2, and although former 7-Mode users will find similarities, major enhancements have been made in this version of SnapVault. One major advance is the ability to preserve storage efficiencies on primary data during SnapVault transfers.

One important architectural change is that SnapVault in clustered Data ONTAP replicates at the volume level as opposed to the qtree level, as in 7-Mode SnapVault. This means that the source of a SnapVault relationship must be a volume, and that volume must replicate to its own volume on the SnapVault secondary.

**Note:** Both primary and secondary must be running clustered Data ONTAP 8.2 or later.

SnapVault users can now take advantage of nondisruptive operations, the cornerstone of clustered Data ONTAP. SnapVault administrators can seamlessly rebalance SnapVault primaries and secondaries for performance or capacity needs, because it is now possible to move SnapVault primary and secondary volumes to different aggregates or nodes within a cluster without disrupting SnapVault operations. If a SnapVault transfer is in progress when a volume is moved by using vol move, the transfer may pause for a few minutes during the volume cutover phase, but the transfer resumes from the most recent transfer checkpoint after the vol move operation completes. Administrators never have to reconfigure a SnapVault relationship just because a volume was moved to another node by using the vol move operation.

A new feature for the clustered Data ONTAP version of SnapVault is the ability to preserve the storage efficiencies of the primary volume throughout a SnapVault transfer. This means that if deduplication and compression are enabled on the primary volume, these efficiencies are preserved during SnapVault transfers, resulting in less data being transmitted over the network and leading to shorter backup windows and bandwidth savings. It also means that the data is already deduplicated and compressed on the secondary volume after the SnapVault transfer completes, without having to run the deduplication and compression processes on the secondary volume.

For the SnapVault transfer to be successful, it needs at least the same amount of free space available on the secondary volume, as if it had to copy the data from the primary volume in a nondeduplicated and noncompressed format. During the SnapVault transfer, it appears that the expanded size of the primary data has been consumed on the secondary. However, as soon as the transfer completes, the storage-efficient dataset size is reflected in the amount of space consumed.

It is possible for deduplication and compression to run on the secondary volume after the SnapVault transfer has completed, independent of the storage efficiencies present on the primary volume. However, enabling the compression process to run on the SnapVault secondary volume causes the storage efficiencies present on the primary volume to not be preserved on the SnapVault secondary volume.

3.4 Sizing Considerations for SnapVault

Correctly sizing a SnapVault solution is an important step in the design phase to make sure that backups complete within a planned backup window, RPOs can be met, and user I/O performance is not adversely affected. There are many variables that must be considered when sizing a SnapVault solution.

For more information (Partners with Field Portal access)... including performance data numbers, please see the ‘Sizing SnapVault’ section in **TR-4183i SnapVault Best Practices Guide for clustered Data ONTAP**.
• **Disk Space:** The SnapVault secondary, or target, must be sized so that adequate disk space is available to retain all of the planned backups. This space can be estimated fairly accurately by using known information about the primary data and the required RPOs. This calculation is independent of the systems used. The data you need to calculate the required disk space on the SnapVault target is the size of the primary data, the daily/weekly/monthly data change rates, the number of daily/weekly/monthly backup copies to be kept, and the space savings that can be anticipated by using NetApp’s deduplication and compression. For information about determining this data and a tool for calculating the space required, see the Rapid SnapVault Space Estimator (Field Portal access required – Internal & Partners only).

• **Data Throughput:** It is also important to size a SnapVault solution so that data can be transferred fast enough to complete backups in the amount of time available. To do this, first determine how much data, on average, is transferred in a SnapVault incremental update and how much time must be allotted to complete the backup, then use these numbers to determine the required data throughput in units such as megabytes per second. If you have multiple SnapVault relationships, remember to consider the size of all the backups that must be completed during a given backup window. Once the required data throughput speed is determined, an appropriate NetApp FAS system can be selected that can handle the data speed required.

• **Storage Efficiencies:** It is important to remember that SnapVault operating in clustered Data ONTAP 8.2 supports storage efficiency preservation during SnapVault transfers. This means that SnapVault initializations and updates are faster if compression and/or deduplication are enabled on the primary data. Again, datapoints concerning these factors can be found in the above linked TR-4183i.

• **Client I/O Impact:** Another factor to consider when sizing a SnapVault solution is the impact of the SnapVault processes on the other workloads running on a system. Clients generally experience more latency with a greater number of concurrent SnapVault streams. If a large number of transfers need to take place during a given backup window, it may be best to stagger the updates so that they don't all run at the same time and cause unacceptable latency to clients. For example, if 80 relationships must update within a 10-hour window and it is determined that, on average, 8 concurrent updates can finish in 1 hour, then a transfer schedule can be created that starts 8 new updates every hour during the given 10-hour window.

When sizing a SnapVault solution, an acceptable client I/O latency threshold should first be established. Then determine what the client I/O latency impact would be for a specific configuration, given the system being used, the type of dataset to be backed up, and the number of concurrent transfers needed to complete all backups within a given window.

### 3.5 Selective Copy

A selective copy allows you to use SnapVault to copy data from a specific source copy. The source copy can be either a primary or a synchronous copy. Selective copy provides the freedom to take frequent SnapShots on Primary, but only vault selective ones (i.e. 1 per day) to a secondary location for backup purposes. This helps control secondary storage usage.

**Note:** Selective copy is applicable for full backups only

You can define a selective copy to be:

- Time-based
For example, only the first or last full backup that occurs within the selected week, month, quarter, half-year, or year will be copied. Note that, during these selection criteria, it considers the Start Time of a backup job.

• Automatically selected
  o For example, if the copy is defined as All Fulls, all full backups from the primary copy will be copied during an Auxiliary Copy operation.

• Automatically not selected on the primary copy
  o You can manually select the jobs to be copied.

3.6 Data Movement to Tape

The SnapProtect solution allows several ways for backup copies to be replicated to tape, including NDMP dump, SMTape, and tape streaming through media agents. Dump is a Snapshot copy based backup and recovery solution where files and directories are backed up to tape and restored to a storage system. SMTape is a Snapshot copy based backup and recovery solution where required metadata and data blocks of a flexvolume is backed up to tape and restored to a storage system. Tape backups are enabled in the storage policy properties under the Snapshot tab; select Enable Backup Copy.

Tape backups can use the Primary(Classic) copy in the storage policy. Alternatively, you can create a new primary backup copy for the storage policy. Tape backups can be configured so that any of the Snapshot copy locations can be used as the source for the tape copy. Select the source for the backup copy in the properties for the storage policy; under the Snapshot tab, modify Source Snap Copy.

The NetApp NAS NDMP iDA is the only agent that allows NDMP dump and SMTape. Tape copies from other iDAs stream through a media agent in a more traditional backup fashion.

Starting with clustered Data ONTAP 8.2, limited SMTape functionality was introduced to allow SnapMirror and SnapVault relationships to be seeded by copying the initial baseline to tape and restoring it to the secondary. Seeding relationships is the only supported use of SMTape in clustered Data ONTAP 8.2.

Creating full or incremental copies to tape depends on how the SnapProtect backup was initiated. If it is required to run incremental tape jobs, then incremental schedules must be created for the SnapProtect backup. To select the tape backup jobs to be moved to tape, right-click the storage policy ➔ Properties ➔ Snapshot tab and modify the Selection Rule.

Starting with 8.2 clustered Data ONTAP, there is support for Storage Virtual Machine aware NDMP backups. This helps in optimizing NDMP backup performance by choosing efficient data transfer paths that are fully compatible with integrated non-disruptive operations and volume mobility capabilities of clustered Data ONTAP.

Cluster Aware Backup (CAB) extensions are required to support these features, and are available in SP2b and newer service packs.

The Clustered Data ONTAP CAB Extensions define a mechanism and protocol within Data ONTAP for establishing efficient data connections for the backup and recovery operations of a storage cluster.

The CAB Extensions present details of the storage cluster’s resource locations as a means to:

• Identify backup resources within the storage cluster:
  o Required backup resources may be located anywhere in the storage cluster
  o Required backup resources may have moved since a previous backup

• To enable a Local backup instead of a 3-way backup when possible

• Prevent data flow across the intra-cluster network

Choosing a Tape Backup Method in 8.2 Data ONTAP 7-Mode

• You should use dump backup and restore if you want the following features:
  a. Individual file and directory backup and recovery
  b. Preserve your backups for several years
c. Exclude specific files and directories during a backup
   • You should use SMTape backup and restore if you want the following features:
     a. A disaster recovery solution that provides high performance
     b. SnapMirror tape seeding
     c. To preserve deduplication and compression during backup and restore operations
     d. Backup and restore of volumes that have millions of files to achieve desired backup windows
     e. Full volume recovery only

Choosing a Tape Backup Method in 8.2 clustered Data ONTAP
   • You should use dump backup and restore if you want the following features:
     a. Preserve your backups for several years
     b. Perform Direct Access Recovery (DAR) of files and directories
   • You should use SMTape backup and restore if you want the following features:
     a. Only SnapMirror Tape seeding from Data ONTAP CLI

For more details on NDMP configuration, see Clustered Data ONTAP 8.2 Data Protection Tape Backup and Recovery Guide.

3.7 Understanding the Backup Workflow of SPOS and clustered Data ONTAP

With the added support for clustered Data ONTAP and limited non-NetApp to NetApp (SPOS), SnapProtect can still manage all backup and recovery workflows of 7-mode, clustered Data ONTAP, and SPOS supported data from a single pane of glass. To understand how the SnapProtect software works, it is important to understand the workflow, starting from the source data and working outward.

Clients own source data. Clients have specific iDataAgents, depending on the type of client and the data being protected. Backup sets and subclients are configured within the iDA, and they group the source data to be protected. For example, if the volume /vol/datavol is to be protected on a client NetApp array, a subclient would contain an entry for /vol/datavol.

Figure 8) SnapProtect Example Workflow: Vault -> Mirror -> Tape.
The storage policy determines the behavior of the data protection operations as well as the retention properties. Each subclient is associated with a storage policy, which contains entries for the various copies in the data protection layout. In the example in Figure 8, client data is protected by NetApp Snapshot copies. Vaulting is performed for longer-term retention. The vaulted data is then mirrored for redundancy. Tape copies are then created from the mirror copy. The SnapProtect software orchestrates the operations, passing the vaulting and mirroring job control to the OnCommand server.

Heterogeneous Support with SnapProtect for Open Systems

SPOS adds support for backup and granular recovery of DAS and non-NetApp storage to NetApp. Its target for backup is Data ONTAP 7-mode SnapVault. SPOS has a block level incremental replication capability, but requires full file system scans. Support includes select non-clustered hosts and applications; these include Windows File System, Linux, Solaris, Microsoft SQL Server, Microsoft Exchange, and Oracle.

SPOS replicates data from Unix and Windows clients with source volumes carved out of disks that are locally attached to the clients or from disks that externally attached to the client from non-Netapp hardware arrays to a Data ONTAP 7-mode destination as a single LUN with a single partition onto a SnapVault Copy. It uses OnCommand Unified Manager to provision the destination volumes to which data is replicated. Data is replicated as GPT LUN on Windows. On UNIX, LUN has the same format as the source volume LUN.

Running SnapProtect for Open Systems jobs from the replication workflow enables the use of a single destination NetApp volume as a fan-in destination for many clients under a single DataSet context. It allows a group of clients to be controlled by a single job management context (providing a many to one control). This group now defines the operations job unit which would trigger the recovery point snapshots on the SnapVault destination.

Using a consolidated dataset with SnapVault automatically pull in any efficiencies that NetApp deduplication can achieve across the collection of the replicated disk images that are maintained with SnapProtect for Open Systems.

Note: A valid SnapProtect controller license is required on the FAS or V-Series controller(s) running Data ONTAP 7-mode, which is the target for SPOS backups. If the target for SPOS is the FAS primary system (which is already licensed as a primary for SnapProtect), then no additional SnapProtect license is required.
Each backup job for a SPOS subclient does the following:

1. Creates a software snapshot of the source data using the native Snap engine such as VSS or Qsnap/LVM.
2. The volume is added to the OCUM dataset that corresponds to the primary snap copy of the storage policy to which the subclient is associated.
3. OnCommand carries out any provisioning tasks and requests the destination filer to connect back to the client SnapProtect for Open Systems Transfer/ SnapProtect for Open Systems Test process.
4. If the volume has never been protected before, then the SnapProtect for Open Systems Transfer will perform a baseline transfer of all the blocks for the volume to the destination filer.
5. If the volume has been protected before, then the SnapProtect for Open Systems Transfer/SnapProtect for Open Systems Test will figure out which blocks have changed for the volume since the last backup and only transfer the changed blocks to the destination file server. At a given volume, all backups, except the first, are incremental.
6. When the SnapProtect for Open Systems transfers to all the associated subclients are complete, a destination snapshot is taken on the destination vault copy and this Snapshot copy is registered as the primary copy of data for that application against the primary snap copy.
7. The software snap that was created at the start of the SnapProtect backup job is deleted once the SnapProtect for Open Systems transfer for source Snapshot copy is successful.
The following illustration depicts the SPOS modules:

Link-1: The backup processes interact with OnCommand Unified Manager server using dfm sdk to add host and volumes to the OnCommand Unified Manager dataset and to run backup.

Link-2: OnCommand Unified Manager server communicates to NDMPListener to get and validate host details.

Link-3: OnCommand Unified Manager server creates/updates relationship with secondary storage and instructs storage to get the data from LREPListener to protect the dataset.

Link-4: When Secondary storage gets a backup request, it connects to the source host on port 10566 to get the replication data from the LREPListener.

The SnapProtect job type maybe full, differential or incremental, but the SPOS transfer is always incremental, once a baseline for a volume is established.

Processes specific to SPOS

- Replication workflow – Central coordinating process which coordinates the backup processes for all the subclients of the replication workflow.
- NDMPListener – Listens on port 10,000. The NDMPListener is required by the OnCommand Unified Manager server for host validation, and file system directory hierarchy discovery and validation. This thread runs on the source host in the context of CVD in a thread and will start with CVD startup.
- LREPListener – Listens on port 10566. It runs on the source host is started on demand when the SnapProtect for Open Systems backup runs. This is the port through which data gets transferred to the destination filer. LREPListener continues to run even after the backup completes.
- LREPCoordinator – Runs on the MediaAgent, associated with the storage policy on which the workflow is triggered. It is responsible for starting the on-demand OnCommand Unified Manager job, which will in turn trigger incremental updates from the source hosts to the destination filer for each of the source volumes involved in the SnapProtect for Open Systems backup job.
Clustered Data ONTAP Support with SnapProtect

SnapProtect for clustered Data ONTAP allows users to perform Snapshot and replication (SnapMirror and SnapVault) in Clustered Data ONTAP environments, from the same interface that they manage their 7-mode, SPOS, etc workflows. A user can protect a clustered Data ONTAP volume by taking its Snapshot and performing replication for the mentioned volume, having a backup copy or a DR copy. The clustered Data ONTAP support is for many protocols (NAS, SAN, NDMP) which makes it possible for users to protect clustered ONTAP volumes, NFS datastores and exports...as well as applications hosted on LUNs.

Storage Virtual Machine (SVM or formerly known as vServer) SnapVault and SnapMirror workflows were added for both Intracluster and Intercluster. SVM aware NDMP backups (support for Cluster Aware Backup extensions) were also added.

Figure 10) SnapProtect for clustered Data ONTAP Workflow.

1. Data is quiesced and protected via application consistent Snapshot copies
2. Snapshot copies and clones are used to access data for indexing (You can also run indexing on the secondary at a later time after backup job has completed using Deferred Indexing – NAS only)
   **Note:** To run NAS Snapshot indexing in SnapProtect for clustered Data ONTAP, use clustered Data ONTAP 8.2P3 or later. NAS Live Browse is not available in clustered Data ONTAP, so if you choose to defer your indexing on the secondary, then you won't be able to browse nor have a catalog of your primary NAS data.
3. OnCommand 6.x handles provisioning of secondary storage, using resource pools and provisioning policies for replication
4. D2D replication - Storage Virtual Machine aware backups using SnapVault / SnapMirror (supports both Intracluster - within the cluster, Intercluster - between the clusters)
5. D2D2T - Storage Virtual Machine aware NDMP backup to tape (support for Cluster Aware Backup Extensions)
In order for SnapProtect to discover volumes on any Storage Virtual Machine, the user will need to add the Cluster Management LIF to the CommCell Console via Array Management.

-Once the Cluster Management LIF is added, SnapProtect can auto-detect SVMs, instead of having to manually add the SVM Management LIF.

Note: From the Storage drop down in OnCommand Unified Manager 6.0 interface, the user is first required to create Resource Pools and SVM Associations (Vserver Associations), the rest can be configured and managed within the SnapProtect interface.

### 3.8 SnapProtect Configuration Overview for clustered Data ONTAP

To initially configure the SnapProtect CommCell for NAS, SAN, and NDMP clients using the Cluster Management LIF:

1. Add the Cluster Management LIF in Array Management. From the CommCell Console, go to Control Panel and select Array Management -> Add -> Enter in LIF IP address. When you enter this Cluster information, upon CommCell Console refresh, the Cluster automatically includes its SVM details and will populate these in SnapProtect's Client section. You can also, alternatively input the SVM (Vserver) LIF information.

<table>
<thead>
<tr>
<th>Cluster Management LIF</th>
<th>SVM (Vserver) Management LIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>You can choose any volume from the Cluster for performing backups.</td>
<td>You can take separate backups of volumes of a particular vserver.</td>
</tr>
<tr>
<td>It allows you to auto detect the existing SVM from Commcell UI.</td>
<td>In order to use SVM IP for configuration, the SVM should have SVM Management LIF.</td>
</tr>
<tr>
<td>You need to provide NDMP credentials for the Cluster while configuration.</td>
<td>You need to provide NDMP credentials for the SVM while configuration of NAS clients.</td>
</tr>
</tbody>
</table>

2. From the same Array Management window, Add the OnCommand server, being sure to select the DataFabric Manager (former nomenclature) check-box.

3. To then configure NAS client:
   a. From the SnapProtect CommCell, right click Client and select All Tasks -> Reconfigure.
   b. Right click the NAS Client, select Properties, click NDMP Properties
      i. Create the NDMP challenge password in the command line interface (CLI) by entering the following command:

```
  vserver services ndmp generate-password -vserver <cluster name / storage virtual machine> -user <Admin/vsadmin>
```

Note: In order to take backup from Cluster Management LIF, node scoped NDMP should be disabled. For more details on NDMP configuration, see [Clustered Data ONTAP 8.2 Data Protection Tape Backup and Recovery Guide](#).
iii. Select Detect to check connectivity with the Cluster and if validated, select OK.

4. To create the Storage Policy:
   a. From the SnapProtect CommCell, right click Client and go to Properties.
   b. Select the Virtual File Server tab, click Detect, and select the Source and Destination SVM.
   c. Click Add and create the Storage Policy in the CommCell (similar to 7-mode Storage Policy creation).

5. Create a Subclient in the SnapProtect CommCell in the same way you would for 7-mode; select the content desired to backup and associate to the Storage Policy created above.

6. You can now perform a backup using this Subclient.

   **Note:** For Application backups, add the SVM for corresponding source and destination volume(s) or LUN(s) in the SnapProtect Console using Auto Detect functionality, and perform backup in the same way as for 7-mode.

### 3.9 OnCommand Utility

When a new replication copy is created in SnapProtect, the task is given to OnCommand for implementation. OnCommand creates the required datasets, provisions the required volumes, and initiates the baseline data transfers. It is important to know how source data gets from a subclient to a dataset to a destination volume during replication. This may differ, depending on the replication type and data type.

In all cases, a SnapProtect subclient has a one-to-one mapping to an OnCommand dataset. In addition, a SnapProtect subclient cannot span clients.

**NAS Data**

For example, if a single NetApp primary system is configured as a NAS client, all of the NAS volumes on that primary system could be grouped together by a single subclient. The result would be a single dataset in OnCommand Unified Manager. If the storage policy in the SnapProtect software calls for mirroring this subclient, then the dataset would create mirror relationships for each of the volumes in the subclient.

In Figure 11, three NAS volumes are grouped into a single subclient. Creating a mirror copy results in a single dataset in the OnCommand server and three mirror relationships are established.

**Figure 11** Mirroring NAS volumes.
Vaulting NAS data is slightly different because a vault can be more granular in scope. In 7-mode, in addition to the entire volume, individual qtrees in a primary volume can be selected for vaulting purposes. In Figure 9, single qtrees from three volumes are grouped into a single subclient. Creating a vault copy results in a single dataset in the OnCommand Unified Manager server and three vault relationships are established, one for each of the qtrees. In this example, OnCommand is configured to allow a fan-in of the vaulted relationships. Enabling fan-in on the OnCommand server requires setting the dpMaxFanInRatio parameter on the server. For example, to set the fan-in ratio to 10, run dfm options set dpMaxFanInRatio=10 on the OnCommand server.

**Note:** A fan-in topology is one where multiple SnapVault primaries back up to one SnapVault secondary. The primary use case for this topology is multiple remote sites that back up to one central data center. In 7-Mode, fan-in meant that multiple qtrees can back up to the same secondary volume. This is referred to as volume-level fan-in. In clustered Data ONTAP, because the relationships are configured at the volume level, it is not possible to back up multiple volumes to one secondary volume. However, it is possible to back up SnapVault primary volumes from multiple Storage Virtual Machines and/or clusters to separate volumes in a single destination SVM, which can be in a different cluster. This is referred to as system-level fan-in.

On the other hand, in 7-mode, if the entire volumes were vaulted, six vault relationships would be established. This is because a relationship for each volume’s non-qtrees data would be created as well as a relationship for each of the volume’s qtrees. In Figure 13, the three volumes are grouped into a single subclient and vaulted (with fan-in enabled).

---

**Figure 12**) Vaulting NAS qtrees - fanin.

**Figure 13**) Vaulting NAS volumes and qtrees - fanin.
LUN Data

When working with LUN data, whether using an application iDA or the file system iDA, there are specific guidelines to follow. Because subclients do not span clients, primary data must be laid out with the subclient in mind.

It is best if all LUNs in the primary volume are protected by a single subclient. For this to work, the following must be true:

- All LUN data on the volume must belong to the same client.
- All LUN data on the volume can be protected by the same iDataAgent.

Volumes with LUN data split across multiple subclients can result in increased capacity requirements for replication operations. Consider an example where three clients (F:\) each map to LUNs in a common volume. Three subclients are used to protect these LUNs. If these subclients were mirrored, it would result in three datasets and three baseline copies for the common volume, as shown in Figure 14. Vaulting would result in the same behavior unless each LUN were in its own qtree. See the resulting wasted capacity with duplicated data!

Figure 14) Poor storage layout with multiple clients with LUN data on a common volume.

See Figure 15 below, the LUNs are provisioned on separate volumes. Notice the difference that results after replication.

Figure 15) Better storage layout with multiple clients having LUN data on separate volumes.
In Figure 16 below, the LUNs on the common volume are mapped by a single client and grouped by a single subclient (rather than three). The result of mirroring this subclient would be a single dataset and a single baseline copy for the common volume. This configuration is ideal to maximize available storage capacity.

Figure 16) Efficient storage layout of single client with LUN data on a common volume.

OnCommand Unified Manager 6.0 Configuration Overview for Clustered Data ONTAP

The OnCommand server configuration is a prerequisite to SnapProtect CommCell configuration.

**Note:** It is recommended that you do not install the CommServe and the OnCommand Unified Manager on the same host.

The following steps describe how to configure OnCommand Unified Manager:

1. Add Cluster
   a. From the OCUM Console, click on Storage and select Clusters
   b. Create and add host information
      i. SVMs (Vservers) will auto populate after Cluster is successfully added

2. Create Resource Pools
   a. From OCUM Console, click on Storage and select Resource Pools
   b. Select Create, and define the Resource Pool(s) name and aggregates you wish you use for your SnapProtect SnapVault/SnapMirror destination target(s).
3. Define the Storage Virtual Machine Association
   a. From the OCUM Console, click on Vserver Associations
      i. In this window, select Create
      ii. Define the source and destination of the SVM and also select SnapMirror or SnapVault in the same window, depending on your use case.

Design Consideration Points:
   i. There are two ways of associating SVMs; you can create an association between any source SVM and destination SVM or between a specified source SVM and a specified destination SVM. Once created, SVMs can’t be modified…you can only delete and re-create a new one.
   ii. By default, SVM associations are not bi-directional. If you have created svmProdLocA -> svmProdLocB association and want to replicate from svmProdLocB -> svmProdLocA, then a new SVM association needs to be created in OCUM for svmProdLocB -> svmProdLocA with the appropriate SnapVault/SnapMirror relationship defined.
   iii. If replicating between SVMs belonging to different clusters (intercluster replication), enable cluster peering (via System Manager or CLI) between the clusters before creating SVM associations between SVMs belonging to those clusters.
   iv. A Storage Virtual Machine association can be created with only one destination SVM per cluster.

   Example: There are 3 SVMs in a 4-Node cluster, svm1, svm2, svm3. If you have already created svm1-> svm2 (SnapMirror) and svm1-> svm3 (SnapVault), OnCommand 6.x will not allow svm1 to associate with any other Storage Virtual Machine within that cluster. However you can still associate svm1 to any other Storage Virtual Machine belonging to a different cluster. However similar limits will apply as shown in this example.

3.10 Replication Options
There are various ways to architect the mirroring and vaulting strategies for replicating data. Mirroring and vaulting can be used separately, or they can be used together and are configured by using storage policies. A storage policy has two copies by default, a Primary(Classic) copy and a Primary(Snap) copy. The Primary(Classic) copy is used for tape copies. The Primary(Snap) copy relates to the NetApp Snapshot copy on the primary system. To vault or mirror that primary data, additional copies must be created in the storage policy.

   -To create a mirror of the primary data, a mirror copy is created that points to Primary(Snap) as its source.
   -Creating a vault of the primary data is similar, also pointing to Primary(Snap) as its source.

To vault the mirror copy, the source for the vault is set to the mirror copy and so forth. Figure 17 shows some of the replication combinations that can be configured.
Note: A fan-out topology is one where a single primary volume is replicated to multiple destinations. This topology allows a single primary volume to be protected and backed up, as well as providing a read-only copy at a secondary site. This is supported in clustered Data ONTAP with a limit of 4 replication destinations for a single source. This limit of 4 destinations is shared between SnapMirror® and SnapVault relationships. There can be any combination of SnapMirror and SnapVault relationships for a single source, as long as the total number of relationships is not greater than 4. Clustered Data ONTAP 8.2.1 supports higher fan-out ratios up to 8.

Scheduling and Retention

Scheduling can be done by creating individual schedules or by creating schedule policies. A schedule policy groups various schedules together, each with its own properties. For example, a schedule policy might contain individual schedules for daily, weekly, and monthly backups. Traditional backup scheduling typically calls for weekly full backups and daily incremental backups. With NetApp Snapshot copy technology; however, the SnapProtect model consists almost entirely of full backups. An exception to this is when protecting NAS data. The indexing performance of NAS data backups increases significantly for incremental backups, while maintaining seamless searching for single file recoveries across Snapshot copies. NAS data on 7-mode also supports the Live Browse feature, where a user can opt out of indexing the contents (i.e. millions of files where indexing would be intensive/lengthy) and can simply browse upon the backups in real time – based on needs.

Auxiliary copies can be scheduled based on specific times, or they can be configured to run automatically.

Figure 18 shows a storage policy and its associated snap copies. In this example, a subclient is scheduled to perform a full backup (a local Snapshot copy) each day at 6:00 p.m. Retention for these local Snapshot copies is configured in the Primary(Snap) copy. A retention model of 10 daily backups and 6 weekly backups is established.

Mirroring takes place each day at 6:30 p.m. Retention for the mirror copy matches that of the primary copy. Vaulting starts when the mirror finishes. Retention for the vault destination is configured in the vault copy. A retention model of 90 daily backups and 52 weekly backups is established.
There are two types of retention rules in the SnapProtect software: basic retention rules and extended retention rules. Basic retention rules apply to daily or hourly backups. Extended retention rules apply to longer-term retention such as weekly full, monthly full, and yearly full backups. These rules are configured in the storage policy and can be set for the primary snapshot copy, vault copies, and tape copies. Mirror copies do not allow specific retention settings, because they inherit the same retention as the primary copy.

A cycle represents a full backup and the incremental backups that depend on that full backup. In many cases, full backups are used for every backup. However, for NAS data with millions of objects, a strategy that includes incremental backups improves indexing performance. In addition, if backup jobs require incremental copies to tape, then the SnapProtect backups on the primary storage must include incremental jobs. Considering a full-backup-only paradigm, each backup can be considered a cycle. When incremental backups are included, all of the Snapshot copies in the cycle are retained until the last incremental in the cycle has expired. Performing more frequent full backups reduces the number of Snapshot copies associated with a cycle.

Basic retention rules allow retention entries for days and cycles. The default setting is 7 days and 2 cycles.

Extended rules can be applied for longer retention. These rules include options to keep all full backups, weekly full backups, monthly full backups, quarterly full backups, half yearly full backups, and yearly full backups. Extended rules are not tied to a particular backup schedule. Rather, they are tied to full backups that start on a particular day of the week or day of the month. These days can be chosen as required.

To perform both hourly backups and daily backups, separate backup sets and storage policies can be created. One backup set includes a subclient with the hourly schedule and associated with one storage policy. The other backup set includes a subclient with the daily schedule and associated with the other storage policy. When running hourly backups, it is necessary to change the data aging schedule to run hourly instead of the default setting of once per day. The data aging operation is what expires backups and deletes Snapshot copies.

**Note:** Many of the application iDataAgents do not allow separate backup sets.

**Enterprise Scheduling Policy Example**

The following examples describe how to keep 6 hourly backups, 30 daily backups, weekly backups for 3 months, and monthly backups for 1 year on the NetApp primary system. These examples assume that only full backup jobs are being scheduled.
• **Hourly backups with 6 hour retention** create a daily schedule for the subclient that repeats every hour, then set a basic retention rule in the primary snap copy of the associated storage policy to retain 0 days and 6 cycles. By default, the data aging schedule runs once per day; therefore, to expire backups based on hourly retention, the data aging schedule needs to run hourly.

• **Daily backups with 30 day retention** create a daily schedule for the subclient that repeats every day, then set a basic retention rule in the primary snap copy of the associated storage policy to retain 30 days and 30 cycles.

• **Retain weekly backups for 3 months** retain one daily backup every week and keeps it for 90 days. This requires an extended retention rule. Set an extended retention rule in the primary snap copy of the associated storage policy to retain weekly full backups for 90 days, and set the rule to start on the appropriate day of the week. Every daily full backup created on this day of the week is retained for 90 days.

• **Retain monthly backups for 1 year** retain monthly backups by retaining one daily backup every month and keeping it for 365 days. This requires an extended retention rule. Set an extended retention rule in the primary snap copy of the associated storage policy to retain monthly full backups for 365 days and set the rule to start on the appropriate day of the month. Every daily full backup created on that day of the month is retained for 365 days.

• **For replication (mirroring and vaulting)**, similar methods can be used to schedule replication and vault retention. However, scheduling is set on the mirror or vault copy in the storage policy rather than the subclient; for vaulting, retention is set in the vault copy in the storage policy.

**Error! Reference source not found.** is an expanded example in which virtual machines (VMs) in VMware® datastores are protected with various requirements. Four different storage policies are required because there are mixed retention requirements across the datastores.

Table 3) Enterprise example of scheduling and retention.

<table>
<thead>
<tr>
<th>Backup Set</th>
<th>Data-Store</th>
<th>Sub-Clien t</th>
<th>Storage Policy</th>
<th>Backup Schedule</th>
<th>Local Retention</th>
<th>Mirror Schedule</th>
<th>Mirror Retentio n</th>
<th>Vault Schedule</th>
<th>Vault Retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>DS1</td>
<td>SC1</td>
<td>SP1</td>
<td>Daily full at 6 p.m.¹</td>
<td>10 days (cycles), 6 weeks, set in primary snap copy</td>
<td>Daily at 6:30 p.m., set in mirror copy schedule</td>
<td>10 days, 6 weeks</td>
<td>After mirror finishes</td>
<td>90 days (cycles), 52 weeks, set in vault copy</td>
</tr>
<tr>
<td>A</td>
<td>DS2</td>
<td>SC2</td>
<td>SP2</td>
<td>Daily full at 6 p.m.¹</td>
<td>30 days (cycles), 8 weeks, set in primary snap copy</td>
<td>Daily at 6:30 p.m., set in mirror copy schedule</td>
<td>30 days, 8 weeks</td>
<td>After mirror finishes</td>
<td>180 days (cycles), 52 weeks, set in vault copy</td>
</tr>
<tr>
<td>A</td>
<td>DS3²</td>
<td>SC3</td>
<td>SP3</td>
<td>Daily full at 6 p.m.¹</td>
<td>10 days (cycles), set in primary snap copy</td>
<td>Daily at 6:30 p.m., set in mirror copy schedule</td>
<td>10 days</td>
<td>After mirror finishes</td>
<td>90 days (cycles), 52 weeks, set in vault copy</td>
</tr>
<tr>
<td>B</td>
<td>DS3²</td>
<td>SC4</td>
<td>SP4</td>
<td>Hourly, except 6 p.m.</td>
<td>23 hours (cycles), set in primary snap copy</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

¹ The local Snapshot schedules can be set at the subclient level or at the backup set level. In the example in Error! Reference source not found., all local Snapshot copies run at 6 p.m.; therefore a single
schedule at the backup set level could be used. If subclients in a backup set require different local
Snapshot schedules, then the schedules need to be set at the subclient level.

2 Datastore DS3 is defined in two backup sets (A and B), because of the need to perform both daily
backups and hourly backups for this datastore. Therefore, it is necessary to have the same datastore in
two subclients to make basic retention rules work. Using two backups sets in this case enables retention
for both hourly and daily backups.

3.11 Understanding the Restore Workflow

With SnapProtect management software, recovery is simple; data can be restored from virtually any
backup copy in a single operation. Restores from recent backups might come from local Snapshot copies
(leveraging optimal Single File Snap Restore technology), while historical data might come from vault or
tape copies.

Data to be restored can be located either by browsing or by using the find feature. When the data is
located, a restore job can be initiated. Restores can be done from any of the backup copies by browsing
data from a particular copy. The copy order precedence is defined on the Copy Precedence tab of the
storage policy properties.

For volumes and LUNs on NetApp primary storage, it is also possible to revert from a Snapshot copy.
This feature uses NetApp SnapRestore® data recovery software to revert a volume or LUN back to a
particular point in time. This feature should be used with caution, because a revert affects all data in the
volume or LUN. To initiate a revert, right-click the subclient and select List Snaps. From the list, right-click
a Snapshot copy and select “Use hardware revert capability if available.”

Because SnapProtect uses NetApp Snapshot technology, it is possible to copy data directly from a
Snapshot copy using CIFS or NFS.

3.12 Data Cloning

SnapProtect software enables administrators to create data clones, which allow read/write access to the
backup data. Data clones can be used for a variety of purposes. NAS data can be cloned when using the
NetApp NAS NDMP iDA. This functionality creates a NetApp FlexClone® volume and makes the entire
contents of the Snapshot copy accessible.

By using a File System iDA, LUN data can be cloned as well. This functionality, when performed on a
primary Snapshot copy, uses LUN clone technology. When cloning LUN data from a Snapshot copy on a
secondary or tertiary NetApp system, a FlexClone volume is created. These FlexClones can be used for
Test/Dev purposes and won’t consume any additional space until written to, and even then….will only be
the deltas occupying additional capacity.

Note: To create data clones, right click the subclient and select List Snaps. From the list, right-click a
Snapshot copy and select Mount..

You can duplicate SQL or Oracle database by cloning the NetApp snapshots created during their
backup. This operation utilizes the cloning capability of the storage array hardware, which enables you to
duplicate large databases within a short period of time. DB clones can be used for multiple purposes,
such as:

- Test environments to troubleshoot issues found in the production database.
- Quick data retrieval without running resource intensive restores on the production environment
  and without requiring additional space on destination server.
- Alleviate load from the production servers for running reports and queries.
A clone can be created from any full backup. During the clone creation, a reservation period can be specified by the user. At the end of the reservation period, the system automatically shuts down the clone database and frees up all resources. The cloning operation can also be scheduled to run periodically to allow the clones to get refreshed from the latest backups on a regular basis.

These SQL and Oracle databases can be cloned to the same instance or to a different instance.

**Note:** DB consistent clones can be done via right clicking on the subclient (SQL or Oracle), digging into All Tasks, and selecting Clone.

### 4 Application Data

SnapProtect can be used to protect applications running on physical servers and hosted on NetApp primary storage. For each supported database application there is an associated iDataAgent. This iDA must be installed on the client system that is running the application. The iDAs prepare the database applications for backup consistency. In addition, they handle things like log truncation during backup, database storage mapping, and log manipulation during restore.

#### 4.1 Support

Figure 19) Application support, iDA, and restore granularity.

#### Protection with application agents

<table>
<thead>
<tr>
<th>Application</th>
<th>iData Agent</th>
<th>Restore Granularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAS</td>
<td>NetApp NAS NDMP iDA</td>
<td>qtrees, directories, files¹</td>
</tr>
<tr>
<td>VMware</td>
<td>Virtual Server iDA</td>
<td>VMs, VMDKs, files²</td>
</tr>
<tr>
<td>Hyper-V</td>
<td>Virtual Server iDA</td>
<td>VMs, files</td>
</tr>
<tr>
<td>SQL Server</td>
<td>SQL Server iDA</td>
<td>Database, filegroup³</td>
</tr>
<tr>
<td>Exchange</td>
<td>Exchange Database iDA</td>
<td>Information Store, DAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Object⁴</td>
</tr>
<tr>
<td>Oracle</td>
<td>Oracle iDA</td>
<td>Database, table</td>
</tr>
<tr>
<td>DB2</td>
<td>DB2 iDA</td>
<td>Database</td>
</tr>
<tr>
<td>SharePoint</td>
<td>SQL iDA, SharePoint Server iDA</td>
<td>DB, Object⁵</td>
</tr>
<tr>
<td>Active Directory</td>
<td>Active Directory iDA</td>
<td>Object⁶</td>
</tr>
<tr>
<td>Lotus Domino</td>
<td>LN DB iDA, LN Document iDA³</td>
<td>DB, DB+trans logs, Doc⁹</td>
</tr>
</tbody>
</table>

#### Protection with virtualization agent

<table>
<thead>
<tr>
<th>Application</th>
<th>iData Agent</th>
<th>Restore Granularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exchange</td>
<td>Virtual Server iDA</td>
<td>VMs, VMDKs, files⁷</td>
</tr>
<tr>
<td>SQL Server</td>
<td>Virtual Server iDA</td>
<td>VMs, VMDKs, files⁷</td>
</tr>
</tbody>
</table>

Footnotes:

1. The NAS iDA can restore qtrees, directories, and files. In addition, it can use SnapRestore on the primary array to revert an entire volume. Live Browse functionality available from Snapshot without indexing (7-mode only).

2. Single file restore (from a NetApp snapshot) for VMs only works for Windows VMs. For Linux, this only works when the backup is to a disk library or tape library (streamed copy). It is not currently possible to
perform granular indexing or single file restore for Linux VMs from the NetApp snapshot using SnapProtect.

3. SQL Server supports point-in-time restore with log replay.

4. Message level restores are available via mining operations. You can use either the offline mining tool or you can configure and perform snap mining.

5. Document level restores are available via snap mining operations.

6. Active Directory backups are accomplished via streaming backup from the Active Directory server. System State backups can also be done via streaming backup using the Windows File System iDA. This requires that you disable system state backup from the default subclient and enable system state backup in a new subclient (with no other content defined).

7. When Exchange or SQL are virtualized and the databases are on VMDKs you can use the virtual server agent (VSA) to backup the VM and perform database consistency during the VM backup. Recovery is the VM, VMDK, or files and not specific to Exchange or SQL. The options for Exchange allow for log truncation during the VM backup. There are no option to truncate SQL logs using this approach. When protecting Exchange in this manor (using VSA), DAG is now supported.

8. LN = Lotus Notes

9. Lotus Notes Database iDataAgent supports both the database and the database plus transaction log restore operations. Lotus Notes Document iDataAgent supports the restoration of documents. Please refer to Books Online for more detail on restore options.

5 Virtualization Data

A key feature of SnapProtect management software is the ability to protect many virtual machines very quickly. In addition, it can index the contents of each VM, and it allows different levels of recoverability, including single file recovery.

SnapProtect software is flexible and allows discovery rules to be established so that new virtual machines can be automatically added to a subclient and protected. For example, using a discovery rule of Datastore Affinity automatically protects new virtual machines on specific datastores.

SnapProtect software uses the Virtual Server Agent (VSA) to perform the data protection operations for virtual environments. The VSA is installed on a system configured as a media agent. Within the VSA, instances are created that define the type of virtualization solution being used. In a VMware environment, a VMware instance would be created under the VSA. Within the instance, a backup set contains the subclients. Figure 20 shows the VSA layout.

Figure 20) Server Agent layers.
**Note:** Because of the advantages of VMware HotAdd transport mode during restores, NetApp recommends installing the VSA on a virtualized media agent. This virtualized media agent should run on an ESX host that has access to the production datastores, such as an ESX proxy host.

In the example shown in Figure 21, multiple datastores are grouped into a single backup set. However, because the datastores have different scheduling and retention requirements, they are separated into their own subclients (each subclient is associated with a different storage policy). The datastores are mirrored and then vaulted.

![Datastores in separate subclients based on differing storage policies.](image)

In Figure 22, the datastores are grouped into a single backup set and a single subclient. In this example, the datastores have the same scheduling and retention requirements. The datastores are mirrored and then vaulted. Because the datastores are grouped into the same subclient, it is possible to do a fan-in on the vault copy.

![Datastores in the same subclient, fan-in.](image)

Backup settings allow different granularity for restore operations. During restore operations, data for the VMs can be browsed and recovered based on the recovery type selected. A container restore can be performed to recover an entire VM. Individual files can also be restored for Windows® VMs, and may leverage SFSR if residing on Primary.

### 5.1 VMware and Applications

When run inside a virtual machine, Microsoft Exchange and SQL Server have integration with VSS that allows database consistency during the backup of the virtual machine. The File System iDA and the VSS
Provider must be installed on the guest OS to get this functionality. To enable these application-consistent backups, make sure that the “Application aware backup for granular recovery” box is checked under the SnapProtect Operations tab for the subclient. Exchange backups offer the additional option to perform log truncation as part of the backup operation; select Truncate ExDB Logs.

Consistent out-of-place restores of SQL Server and Exchange databases can be performed by restoring the flat database files. The Exchange Offline Mining tool is a standalone utility that allows individual message restores from a backup copy of the Exchange database.

6 Summary

NetApp's SnapProtect management software is changing today's backup and recovery landscape. SnapProtect software combines simplified manageability, power, and flexibility for virtual environments with full support for enterprise database applications while providing virtually seamless integration with NetApp Snapshot® technology for fast and efficient backup operations. In addition, SnapProtect software integrates with NetApp SnapVault and SnapMirror software with support for content-based cataloging and movement to tape-based media. This document covered an introduction to the SnapProtect solution along with a design overview. It provided an overview of the technology, and describes some of the basic configuration steps required to get started. SnapProtect offers single pane of glass management for backup and recovery workflows of NetApp 7-mode, clustered Data ONTAP, limited non-NetApp to NetApp (SPOS), and much more. Centralizing all of these functions, along with with policy based management and granular recovery across all supported workloads, makes for a very compelling Enterprise Backup and Recovery solution.
Resources

For more detailed installation information refer to the SnapProtect Online Documentation (Books Online).
http://support.netapp.com/NOW/knowledge/docs/snapprotect/relsnap_protect100sp4/215-08433_A0_books_online_100sp4/books_online_1/default.htm

NetApp partners can find additional SnapProtect information, ranging from Training Presentations, Best Practice Guides, Cookbooks, Datasheets, Technical Reports, to a FAQ doc, on the Field Portal.
https://fieldportal.netapp.com/fas/snapprotect.aspx#15484

A consolidated repository, access and reporting tool for NetApp’s product compatibility information.
http://support.netapp.com/matrix